

A Role for Biotechnologies in the Improvement of Oil Palm Sustainability

Alain RIVAL and Estelle JALIGOT

Cirad Biological Systems

UMR DIAPC - Palm Development Group

34394 Montpellier Cedex5,

France.

Introduction

Biotechnologies and sustainable development are often considered as antagonists, especially in Europe where GMOs (Genetically Modified Organisms) and agricultural biotechnology are opposed to more environment-friendly agricultural practices, such as organic farming.

Nevertheless, there is increasing evidence for a pivotal role for this approach in the ecological intensification of oil palm cultivation. Ecological intensification is based on the understanding of biological functions in order to better exploit natural living resources while preserving them for future generations. Biotechnology is based on mechanisms that nature already uses. When applied in appropriate ways the results can offer new solutions to existing problems, such as the finite resource levels being depleted today (Mc Laren, 2005).

Modern biology is now integrating systems biology with very close links with biotechnology. Indeed, living organisms are supported by the genome (DNA) through the action of the transcriptome (RNAs), proteome, metabolome, and ionome, the four basic pillars of functional genomics (Figure1). These pillars represent the sum of all the expressed genes, proteins, metabolites, and elements within an organism. The dynamic response and interaction of these biochemical “omes” defines how a living system functions, and its study, systems biology, is now one of the biggest challenges in life sciences.

In oil palm as in many major crops, functional genomics is still at its infancy, although there are no reasons why oil palm should not rapidly benefit from the fast progresses generated by automated and high-throughput technologies.

To date, 1193 genomes have been fully sequenced, including 77 archaeal, 996 bacterial and 123 eukaryal genomes from which ten are from higher plants, namely: *Cucumis*, *Vitis*, *Populus*, *Oryza*, *Arabidopsis*, *Glycine*, *Physcomitrella*, *Sorghum*, *Vitis* and *Zea* (GOLD, 2010).

The recent success of sequencing projects on model plants has created a widespread interest in exploring the structure and expression patterns of the genome in several crop species of agronomic interest, and many more genome sequences are to be released in the coming months for plants such as cassava, tomato, potato, papaya, cotton, eucalyptus, citrus, barley,...(NCBI, 2010).

Several institutions and consortia have now achieved the full sequencing of the oil palm genome, paving the way for rapid changes in many genomics-based approaches.

Oil palm breeding has provided an average 1% genetic gain per year since the early 60s and such an impressive increase in oil yield will be maintained in future generations with a major contribution from biotechnology. Indeed, the recent adoption of biotechnological approaches for molecular breeding has already proven very useful in major areas such as the cloning of outstanding plant material, the identity checking of progenies/mother palms or the identification and characterisation of genes underlying useful agricultural traits.

Phenotypic differences among individuals are partly the result of quantitative differences in transcript abundance. Using microarray technology, it is now possible to assess the abundance of many transcripts - and, indeed, of the entire known transcriptome - simultaneously. Studies which attempt to localize the genetic regulators of gene expression have been carried out in several crop species, including maize and eucalyptus (Figure 2). Microarrays are used to determine gene expression levels in segregating populations and identify genomic regions (gene expression QTLs, or eQTLs) explaining transcript variation in co-regulated genes.

The present article will focus on recent developments in various plant biotechnology areas in which applications for improving oil palm sustainability already exist or will be developed in the very near future.

Intensification is no more an option for oil palm

To meet the food and bio energy security needs of the near future, farmers must double or even triple crop yields on less arable land, with less water, on poorer soils, and with less fertilisers and pesticides. Information gained from sequenced genomes in crops, coupled with genetic association studies, will allow us to identify key genes/quantitative trait loci and networks that can lead to higher yielding crops that can grow in extreme conditions with reduced environmental impact.

The oil palm being a major crop in the intertropical region -where fast growing populations live and where large areas of HCV (High Conservation Value) native forests must be preserved- must follow and even anticipate this trend towards ecological intensification (Fitzherbert et al, 2008; Koh and Wilcove, 2008; Stone, 2007).

Given the insatiable demand to supply the expanding population of India, China and Pakistan, and their concomitant increase in standard of living, Murphy (2007) identified the yield problem as one of the major challenges the oil palm industry will have to tackle in the 21st century.

Deciphering the oil palm genome has become reality

Plant genome sequencing has progressed rapidly since the first genome (*Arabidopsis thaliana*) was completed in 2000 (Arabidopsis Genome Initiative, 2000). Then the 389-Mb rice genome was completed in 2004 (International Rice Genome Sequencing Project, 2005), and a sequence of the 2,3-Gb maize genome was recently released (Wei et al, 2009).

All were sequenced using “traditional” sequencing approaches in which sequencing libraries are constructed from individual segments of the genome (such as those contained within bacterial artificial chromosomes - BAC clones) and are sequenced via gel electrophoresis (Sanger sequencing).

A whole-genome shotgun (WGS) strategy, made possible with improved assembly algorithms, has been used for several recent plant genomes, like poplar (Tuskan et al, 2006), grapevine (Jaillon et al, 2007) or sorghum (Paterson et al. 2009) in which the sequencing libraries are made directly from genomic DNA. These approaches are easier and cheaper than the Sanger-based ones.

Recently, next generation sequencing (NGS) technologies have promised to further accelerate progress, with huge increases in sequencing throughput and decreases in costs. Indeed, this approach was found able to generate the genome of all 12 chromosomes (~400 Mb) of *Oryza barthii* in a month or two (Rounsley et al, 2009).

Four different NGS technologies are now commercially available, namely: 454 Life Sciences (acquired by Roche), Solexa (acquired by Illumina), ABI SOLID (acquired from Agencourt Biosciences), and Helicos Biosciences. Although all have their specific features, generally, they can be grouped into two classes based on the lengths of the sequence reads which are generated. Indeed, Solexa, ABI SOLID, and Helicos all produce very short reads in very large quantities, while the 454 platform can produce a more moderate amount of sequence, but with much longer read lengths. Several of the platforms have already gone through multiple rounds of upgraded specifications, and improvements are likely to continue.

Future improvements in NGS technologies may continue to reduce the time and cost of such projects, but this strategy will likely always be valuable for addressing the complexities of sequencing large repetitive genomes with NGS. It provides an immediate and practical solution to the rapid generation of genome sequences from large and complex eukaryotic genomes like the oil palm (1.7 Gb; Rival et al, 1997).

Several research groups and consortia have recently announced the complete sequencing of the oil palm genome. Indeed, Zieler et al (2010) very recently reported on the generation whole-genome

shotgun sequences of the oil palm genome, using primarily Sanger reads to enable high-quality assemblies. This breakthrough was achieved by a consortium joining Synthetic Genomics Inc. (USA), ACGT Sdn. Bhd. (Malaysia) and the J. Craig Venter Institute (USA). According to these authors, the genome sequences have been supplemented with a high volume of EST and transcriptome sequencing and the genome has been fully annotated.

In November 2009, The Malaysian Palm Oil Board (MPOB) announced that a consortium co-led by the Advanced Biotechnology and Breeding Centre has sequenced three oil palm genomes from the two species *Elaeis oleifera* and *Elaeis guineensis*, including the *pisifera* and *dura* palms. The consortium included St. Louis, Missouri-based Orion Genomics, MOgene LC and The Genome Center at Washington University, South Korea-based Macrogen Inc., and Adelaide, Australia-based GeneWorks Pty Ltd. In addition to sequencing and assembling the genomes of the three oil palm varieties, the consortium sequenced the expressed genes (or transcriptome) from multiple tissue types for all three types of oil palm (Wahid, 2009).

In May 2009, a Malaysian consortium created by Sime Darby Bhd, one of the world's largest oil palm company and Synmatix Sdn Bhd, a bio-informatics company had announced the successful sequencing, assembling and annotating of the oil palm genome with 93.8% completeness.

For oil palm biotechnologists all around the world, given the amounts of investments and subsequent issues on intellectual property rights, the availability of these genome sequences to the scientific community still remains a key point. Successful bioinformatics relies on the accuracy and public availability of genome databases, and the oil palm community must follow the example of consortia (public or private) which have already published complete plant genome sequences and permitted their open public access (see the maizesequence.org site as an example among others). Similarly to the "Open Source" concept in software science, each member of the community benefits from the work done by other members, as any improvement made on the published sequence, such as corrections and annotations, is published and made available for further uses.

Biotechnologies already impacts oil palm productivity

The assembly of genomic tools for use in oil palm breeding is clearly under way and it has already delivered outstanding outputs.

Recent review articles published by Price et al (2007) and Rival et al (2007) provided updated information on advances in biotechnology and molecular breeding applications for the oil palm.

The genetic basis of some of the key traits in oil palm (growth, bunch size, bunch number, Oil Extraction Rate) still needs to be dissected with the help of accurate genetic maps. To undertake such a task, molecular resources do exist although they are scattered among various institutions around the world and rarely open to public access. An exception is the LINK2PALM program (<http://www.neiker.net/link2palm/OilP/DefOIL.htm>) which provided valuable resources to the community under the form of co dominant microsatellite markers (SSR).

Microarray (DNA chips) analysis approaches for transcriptomics allows an entire transcriptome to be analysed at a specific developmental point, under a particular stimulus, or in a particular tissue. The construction of such arrays is usually based on very extensive EST collections, if not complete genome sequences. Jouannic et al. (2005) generated five different cDNA libraries from oil palm male and female inflorescence shoot apices and zygotic embryos. A total of 2,411 valid EST sequences were thus obtained through unidirectional systematic sequencing. This EST database described is a first step towards gene discovery and cDNA array-based expression analysis in oil palm.

Such resources are now being generated at a large scale in oil palm as part of an industrial consortium called OPGP for Oil Palm Genome Project, although a limited public access to these resources is anticipated (Mayes et al, 2009).

Recent advances have been made in the exploration of epigenetic mechanisms underlying the *mantled* somaclonal variation in clonally micropropagated oil palms. Indeed, Rival et al (2008) showed that the genome-wide hypomethylation previously described in *mantled* material was not explained by a decrease in expression levels of the *de novo* or maintenance DNA Methyltransferases, a paradox which

has been previously reported in tumour cells. Research is now focusing on DNA methylation around candidate marker genes. Indeed, orthologues of the MADS-box genes involved in the formation of floral organs have been identified in the genome of oil palm (Adam et al, 2006; Syed Alwee et al, 2006), and research has shown that oil palm B-type MADS-box genes display differential transcript levels between normal and abnormal inflorescence tissues (Adam et al., 2007). A range of genes with altered expression in abnormal tissues has been recently identified through the use of subtractive PCR (SSH) and subsequent macroarray hybridization (Beule et al, unpublished data). Investigating methylation patterns of these target genes will pave the way for understanding the epigenetic mechanisms underlying the induction and maintenance of somaclonal variation in oil palm regenerants.

Since the initial transformation of oil palm with a marker gene was been reported by Parveez et al (2000), genetic engineering in oil palm is progressing at the slow pace imposed by public acceptance, biosafety issues and the genetic complexity of desired traits (from the production of bio plastics to the introduction of resistance to major pests and diseases). Subhi et al (2010) have recently identified the ubiquitin extension protein (uep1) gene as a constitutively expressed gene in oil palm. The 5' region of uep1 functions as a constitutive promoter in oil palm and can drive GUS expression in all tissues tested and also be used in dicot systems.

Oil palm physiology and nutrition benefits from biotechnologies

Within the context of global warming, it is of paramount importance to understand the mechanisms underlying the climatic plasticity of oil palm and thus its capacity to withstand dryer climates and produce oil on poorer soils and/or with lower use of fertilizers (Zechendorf, 2009).

In order to study the dynamics of plant metabolism under stress and unravel regulatory mechanisms in place, it is important to combine the traditional more descriptive physiological approaches with the techniques of functional genomics, namely the high throughput methods for transcriptomic, proteomic, metabolomic and ionic analysis. Using this integrated analysis it would be possible to study the dynamics of plant metabolism in the context of the plant system as a whole (Chaves et al, 2009).

Recent advances have been made in the understanding and further modelling of oil palm productivity through the analysis of phenological and growth responses to seasonal and inter-annual climatic variability (Legros et al, 2009). Now is the time to integrate the activity of specific key-genes in these models.

In developing resistance to biotic stresses, the main problem lies in dealing with the highly variable nature of biological agents, be it pathogen or insect. Predicting which genes would confer durable resistance remains a key challenge (Leung, 2008). For abiotic stresses, a high degree of genotype \times environment ($G \times E$) interaction makes assessing the causal relationship between genotype and phenotype difficult. Thus, the challenge is to define phenotypes that are a true reflection of the genotypic differences, and to find the right genes/phenotypes that work well in target environments. This difficulty is exemplified in studying traits responsible for tolerance to drought.

A number of selection criteria have been developed for fast reliable screening of water-use efficiency (WUE) and its determinants such as stomatal conductance and $^{13}\text{C}/^{12}\text{C}$ isotope discrimination. This approach recently proved useful in the understanding of carbon use during the passage from heterotrophy to autotrophy in oil palm leaves (Lamade et al, 2009). Interestingly, a molecular linkage map with several major QTLs for $^{13}\text{C}/^{12}\text{C}$ isotope discrimination in rice was published by Price et al (2002), so any QTL candidate gene regions identified in related species could be potentially anchored to the rice genetic map to provide comparative information on possible homologue location in the oil palm.

Pennisi (2008) published a short review of research under way to identify genes underlying drought resistance in plants. When expression patterns of 1,500 genes were studied in *Arabidopsis* plants grown under drought conditions, 40 appeared to be involved in drought adaptation. Thus there is a palette of genes from which breeders and crop scientists will select for drought tolerance.

The doubling of agricultural food production worldwide over the past four decades has been associated with a 7-fold increase in the use of nitrogen (N) fertilizers. As a consequence, both the recent and future intensification of the use of N fertilizers in agriculture already has and will continue to have major detrimental impacts on the diversity and functioning of the non-agricultural neighbouring bacterial, animal, and plant ecosystems. In a recent review, Hirel et al (2007) described how our understanding of the physiological and molecular controls of Nitrogen assimilation under varying environmental conditions in crops has been improved through the use of combined approaches, mainly based on whole-plant physiology, quantitative genetics, and forward and reverse genetics approaches.

It is of major importance to identify the critical steps controlling plant N use efficiency (NUE). Moll et al. (1982) defined NUE as being the yield per unit of available N in the soil (including the residual N present in the soil and the fertilizer). This NUE can be divided into two processes: uptake efficiency (NupE; the ability of the plant to remove N from the soil as nitrate and ammonium ions) and the utilization efficiency (NutE; the ability to use N to produce yield). Studies of the whole-plant N response will be essential to elucidate the regulation of NUE and to provide selection criteria for breeders and monitoring tools for planters in the aim of designing efficient and reasoned fertilization protocols.

Results from *Arabidopsis* research have improved our understanding of the relationship between N availability, N uptake, and root development (Walch-Liu et al., 2006, Hirel et al, 2007). Since N uptake is one of the most critical NUE components under N-limiting conditions in a number of crops, the transfer of knowledge to crop plants such as oil palms is more than realistic.

Oil palm ionomics is round the corner

Lahner et al (2003) described the ionome to include all the metals, metalloids, and non-metals present in an organism, extending the term metallome to cover biologically significant non-metals such as nitrogen, phosphorus, sulfur, selenium, chlorine, and iodine.

Because the ionome is involved in such a broad range of important biological phenomena, including electrophysiology, signaling, enzymology, osmoregulation, and transport, its study promises to yield new and significant biological insight (Salt, 2004). An understanding of the ionome and how it interacts with other cellular systems such as the genome, proteome, metabolome, and environment are integral to the full understanding of how plants integrate their organic and inorganic metabolisms (Figure 3).

Several recent papers (Baxter, 2009) have described high-throughput elemental profiling studies of how the ionome responds to the environment or explored the genetics that control the ionome. When combined with new genotyping technologies, ionomics provides a rapid way to identify genes that control elemental accumulation in plants.

Inductively coupled plasma spectroscopy, either mass spectroscopy (ICP-MS) or optical emission spectroscopy (ICP-OES) allows for the simultaneous measurement of dozens of elements. Improvements in these techniques over the last few decades have overcome obstacles that prevented previous efforts from detecting all but the most severe ionic differences.

Measuring multiple elements allows researchers to explore the dynamics of the ionome as a whole, not just individual elements in isolation. It also makes gene identification experiments more efficient, which should allow new genetic mapping techniques to identify hundreds of new loci that control this complex system. To fully understand the complex regulation of the ionome, we will need to find the genes that control the accumulation and distribution of each element.

Iron chlorosis is one of the major abiotic stresses affecting crops in calcareous soils and leads to a reduction in growth and yield. In a recent study, Forner-Giner et al (2009) have examined the differential gene expression induced by iron deficiency in the susceptible citrus rootstock *Poncirus trifoliata* (L.) Raf. The genes identified are putatively involved in cell wall modification, in determining photosynthesis rate and chlorophyll content, and reducing oxidative stress.

From Molecular Biology to bioinformatics ... and back

The challenge is to accelerate biological discovery to address some of our most critical food and bioenergy challenges. With the recent sequencing many plant species including major crops, the development of both physical (*in vivo*) and informational (*in silico*) resources has entered a new phase. Bioinformatics has emerged a new field of science aimed at: i) acting as a repository of information generated and to provide links to the associated physical resources the information is based on, ii) To allow the repository to be searched for specific resources or information. iii) To develop linkages between different types of information (sequences, map positions, alleles within a population, related genes) and develop tools to facilitate discovery of further links (gene prediction, protein interactions etc.). To this end, many tools have been developed to generate and improve databases in order to provide simple access and searching facilities and tools for gene discovery. The sophistication of these interfaces continues to improve (Mayes et al, 2005). In the short term, the quantity of data is likely to continue to expand exponentially and ways need to be found to interrogate such databases intelligently, so that the output remains comprehensive.

The present paper was aimed at stressing the importance of biotechnology in the global effort for increasing and maintaining sustainability for the oil palm. For the (still small) community of oil palm biotechnologists, the challenge is very exciting as amazingly powerful tools become available every day. Bioinformatics is emerging as a new science dedicated to the sorting and assembling of DNA and RNA sequences. This *in silico* work is indispensable, given the exponential amount of data accumulated today by high throughput sequencing projects.

If biotechnologists want to play their crucial part in the sustainability challenge for oil palm, they will have to integrate bioinformatics data to biological questions arising from field experiments, in order to provide adequate answers to end-users: the oil palm farmers from tropical regions of the world.

References

- Adam H, Jouannic S, Morcillo F, Richaud F, Duval Y, Tregear JW, 2006. MADS box genes in oil palm (*Elaeis guineensis*): patterns in the evolution of the SQUAMOSA, DEFICIENS, GLOBOSA, AGAMOUS, and SEPALLATA subfamilies. *Journal of Molecular Evolution* 62, 15–31.
- Adam H, Jouannic S, Orioux Y, Morcillo F, Richaud F, Duval Y, Tregear JW, 2007. Functional characterization of MADS box genes involved in the determination of oil palm flower structure. *Journal Experimental Botany* 58, 1245–1259.
- Arabidopsis Genome Initiative, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408(6814):796–815.
- Baxter I, 2009. Ionomics: studying the social network of mineral nutrients. *Current Opinion in Plant Biology*, 12:381–386.
- Chaves MM, Flexas J, Pinheiro C, 2009. Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany* 103: 551–560.
- Fitzherbert EB, Strubbig MJ, Morel A, Danielsen F, CA. Bruhl, PF Donald, Phalan B, 2008. How will oil palm expansion affect biodiversity? *Trends in Ecology and Evolution*, 23(10): 538-545
- Forner-Giner MA, Llosa MJ, Carrasco JL, Perez-Amador MA, Navarro L, Ancillo G, 2009. Differential gene expression analysis provides new insights into the molecular basis of iron deficiency stress response in the citrus rootstock *Poncirus trifoliata* (L.) Raf. *Journal of Experimental Botany*, doi:10.1093/jxb/erp328
- GOLD, Genome On Line Database V3.0, 2010. <http://www.genomesonline.org/gold.cgi>
- Hirel B, Le Gouis J, Ney B, Gallais A, 2007. The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *Journal of Experimental Botany*, 1-19.
- International Rice Genome Sequencing Project, 2005. The map-based sequence of the rice genome. *Nature* 436(7052):793–800.
- Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al, 2007 The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449(7161):463–7.
- Koh LP, Wilcove DS, 2008. Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letters*, 20:1-5.
- Kirst M, Myburg AA, De Leon JPG., Kirst ME, Scott J, Sederoff R, 2004. Coordinated genetic regulation of growth and lignin revealed by Quantitative Trait Locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*. *Plant Physiology*, 135: 2368–2378.
- Lahner B, Gong J, Mahmoudian M, Smith EL, Abid KB, Rogers EE et al, 2003. Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nature Biotechnology* 21: 1215–1221.
- Lamade E, Setiyo IR, Girard S, Ghashghaie J, 2009. Changes in ¹³C/¹²C of oil palm leaves to understand carbon use during their passage from heterotrophy to autotrophy. *Rapid Commun. Mass Spectrom.* 23: 2586–2596.
- Legros S, Mialet-Serra I, Caliman JP, Siregar FA, Clément-Vidal A, Fabre D, Dingkuhn M, 2009. Phenology and growth adjustments of oil palm (*Elaeis guineensis*) to photoperiod and climate variability. *Annals of botany*, 104 (6): 1171-1182.
- Leung H, 2008. Stressed genomics — bringing relief to rice fields. *Current Opinion in Plant Biology*, 11:201–208.
- Mayes S, Parsley K, Sylvester-Bradley R, May S, Foulkes J, 2005. Integrating genetic information into plant breeding programmes: how will we produce varieties from molecular variation, using bioinformatics? *Annals of Applied Biology*, 146:223–237.
- Mayes S, Stadler F, Basu S, Massawe F, Bonin A, Chin CF et al, 2009. Marker assisted selection in plant breeding – generating large datasets for trait analysis. *PIPOC 2009 Int. P.O. Cong. Agric., Biotech. & Sustain. Conf.-Vol. 1*, pp 226-235.
- McLaren JS, 2005. Crop biotechnology provides an opportunity to develop a sustainable future. *TRENDS in Biotechnology*, 23 :7, 339-342.

- Moll RH, Kamprath EJ, Jackson WA, 1982. Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *Agronomy Journal* 74: 562–564.
- Murphy DJ, 2007. Future prospects for oil palm in the 21st century: Biological and related challenges. *Eur. J. Lipid Sci. Technol.* 109 (2007) 296–306.
- NCBI, National Center for Biotechnology Information, 2010. http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList_hn.html#EST
- Parveez GKA, Masri MM, Zainal A, Majid NA, Masani Mat, YunusA, Fadilah HH, et al, 2000. Transgenic oil palm: production and projection. *Biochem Soc Trans* 28:969–972
- Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, et al, 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 457:551–6.
- Pennisi E, 2008. Plant Genetics: The blue revolution, drop by drop, gene by gene. *Science*: 320 (5873): 171-173.
- Price A H, Cairns J E, Horton P, Jones H G, Griffiths H, 2002. Linking drought-resistance mechanisms to drought avoidance in upland rice using a QTL approach: progress and opportunities to integrate stomatal and mesophyll responses. *Journal of Experimental Botany* 53:989–1004.
- Price Z , Mayes S, Billotte N, Hafeez F, Dumortier F , MacDonald D , 2007. Oil Palm, in C. Kole (Ed.) *Genome Mapping and Molecular Breeding in Plants*, Volume 6 Technical Crops, Springer-Verlag Berlin Heidelberg, pp 93-108.
- Rival A, 2007. Oil palm. In: by EC Pua and MR Davey (eds.). *Biotechnology in agriculture and forestry* (vol. 61). Berlin : Springer-Verlag, p. 59-80.
- Rival A, Beulé T, Barre P, Duval Y, Hamon S & Noirot M, 1997. Comparative flow cytometric estimation of nuclear DNA content in embryogenic calli and seed-derived oil palm (*Elaeis guineensis* Jacq). *Plant Cell Reports*. 16: 884-887.
- Rounsley S, Marri PR, Yu Y , He R & Sisneros N, Goicoechea JL et al, 2009. De novo Next Generation Sequencing of plant genomes. *Rice*, 2:35–43.
- Ruffel S, Krouk G, Coruzzi GM, 2009. A systems view of responses to nutritional cues in *Arabidopsis*: towards a Paradigm shift for predictive network modelling, 2009. *Plant Physiology Preview*. DOI:10.1104/pp.109.148502
- Salt DE, 2004. Update on Plant Ionomics. *Plant Physiology*:136: 2451–2456.
- Stone R, 2007. Can palm oil plantations come clean? *Science*, 317:1491-1491.
- Subhi SM, GKA Parveez, I Ismail, 2010. Isolation and characterization of oil palm constitutive promoter derived from ubiquitin extension protein (uep1) gene, *New Biotechnol.* doi:10.1016/j.nbt.2010.01.337
- Syed Alwee S, Van der Linden CG, Van der Schoot J, de Folter S, Angenent GC, Cheah SC et al, 2006. Characterization of oil palm MADS box genes in relation to the mantled flower abnormality. *Plant Cell, Tissue and Organ Culture* 85, 331–344.
- Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al, 2006 The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313(5793):1596–604.
- Wahid MB, 2009. Sequencing the oil palm genome: The beginning. *PIPOC 2009 Int. P.O. Cong. Agric., Biotech. & Sustain. Conf.-Vol. 1*, pp3.
- Walch-Liu P, Liu LH, Remans T, Tetser M, Forde BG, 2006 Evidence that L-glutamate can act as an exogenous signal to modulate root growth and branching in *Arabidopsis thaliana*. *Plant and Cell Physiology* 47, 1045–1057.
- Wei F, Zhang J, Zhou S , He R , Schaeffer M, Collura, K et al, 2009. The physical and genetic framework of the Maize B73 Genome. *PLoS Genetics*: 5 (11): e1000715
- Zechendorf B, 2009. Sustainable development: how can biotechnology contribute? *TIBTECH* 17 : 219-225.
- Zieler H, Richardson T, Schwartz A, Herrgard M , Lomelin D, Mathur E et al 2010. Whole-Genome shotgun sequencing of the oil palm and *Jatropha* genomes. *Plant & Animal Genomes XVIII Conference*, San Diego, CA, USA. January 9-13.

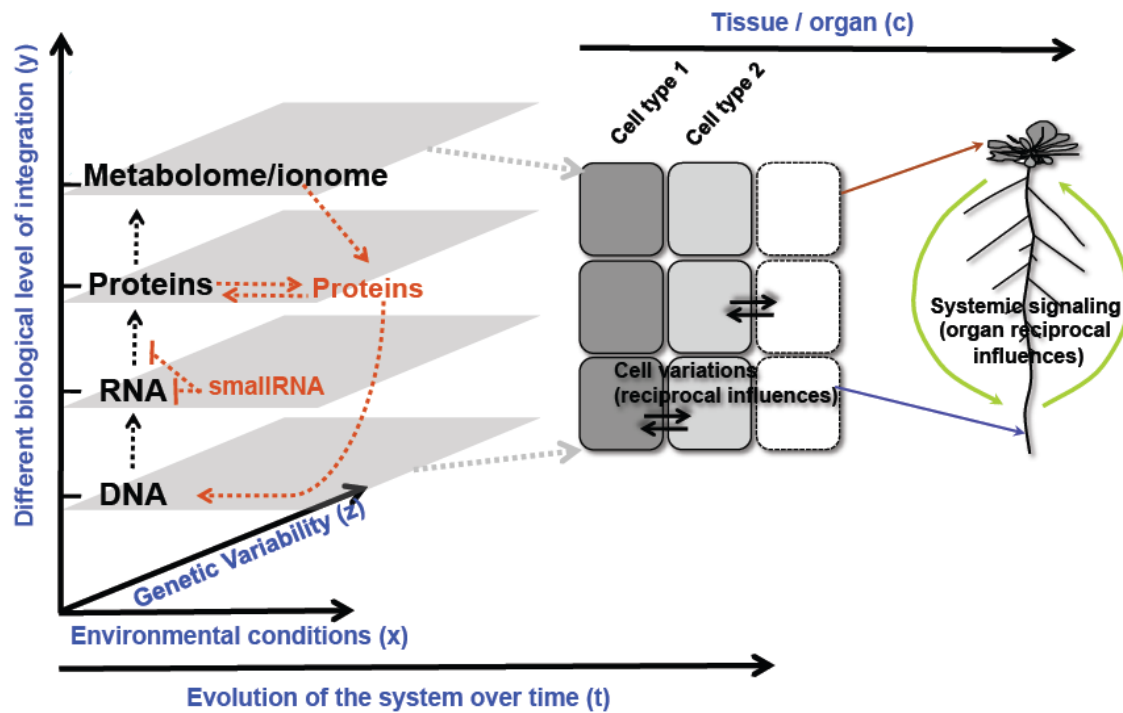


Figure 1 - Plant systems biology: A dimension-wise challenge (Ruffel et al, 2009). This scheme illustrates the different (non exhaustive) dimensions (in blue) that can be integrated in systems approaches, in order to decipher emerging properties of plant regulatory networks.

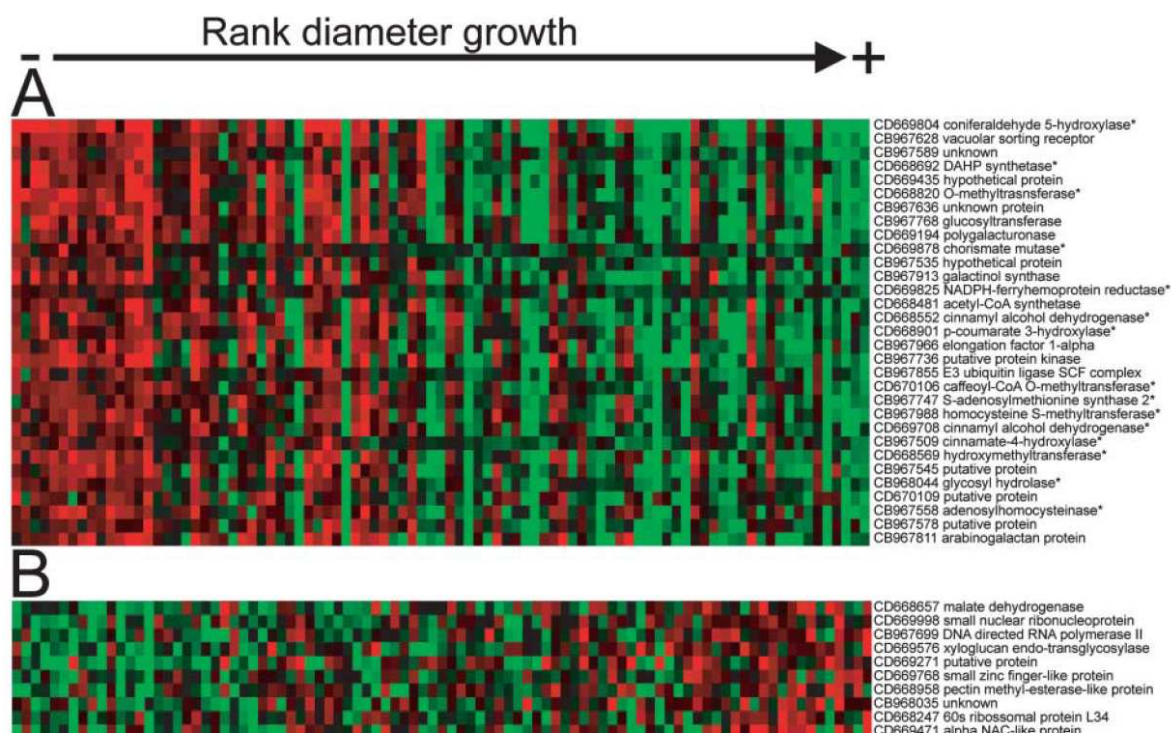


Figure 2 - Transcriptome analysis (Kirst et al, 2004): Association between gene expression and diameter variation. The *Eucalyptus grandis* backcross progeny is ranked according to diameter (x axis) and negative (A) or positive (B) correlation between relative transcript level and diameter variation (y axis). Least square means estimates of transcript levels were normalized relative to the mean of each gene across the population. Red represents higher and green represents lower expression relative to the other individuals of the population, for each specific gene.

Black indicates no change in mRNA levels. GenBank accession numbers and putative functions are displayed on the right. Genes represented by multiple cDNAs are represented by the most highly correlated, and those related to lignin biosynthesis are indicated by asterisks.

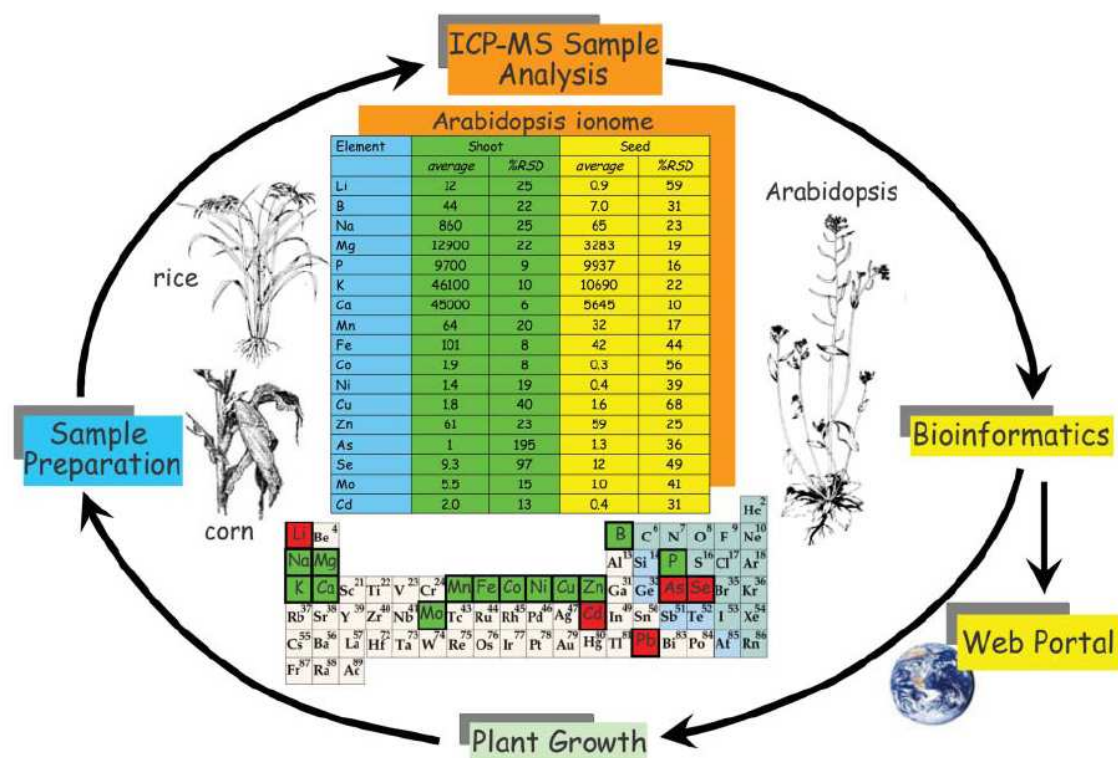


Figure 3 - High-throughput ionomics (Salt, 2004): Putative mutants and wild-type *Arabidopsis* plants are grown together with known ionic mutants, used as positive controls, under standardized conditions. Plants are uniformly sampled, digested in concentrated nitric acid, diluted, and analyzed for numerous elements using ICP-MS. Raw ICP-MS data are normalized using analytical standards and calculated weights based on wild-type plants (Lahner et al., 2003). Data are processed using custom tools and stored in a searchable, World WideWeb-accessible database. Ionomics analysis can also be applied to other plants with available genetic resources, including rice and maize. Elements in the Periodic Table highlighted in black boxes represent those elements analyzed during our ionic analyses using inductively coupled plasma mass spectroscopy (ICP-MS), elements highlighted in green are essential for plant growth, and those in red represent nonessential trace elements. The table represents *Arabidopsis* (Col 0) shoot and seed ionomes, all elements presented as mg/g dry weight. Data represent the average shoot concentrations from 60 individual plants and seed from 12 individuals 6 SD as percentage of average (%RSD).

[Draft and Slide Show]



International Conference on Oil Palm & Environment

Measurement and mitigation
of environmental impact of palm oil production

23 – 25 February 2010 • Grand Hyatt Hotel, Nusa Dua Bali - Indonesia

A Role for Biotechnologies in the Improvement of Oil Palm Sustainability

Alain RIVAL, Estelle JALIGOT
Cirad BioS - Palm Development Group
Montpellier, France

Acknowledged by:



Organized By:



WWW.ICOPE-SERIES.com

Sustainable Development and Biotechnologies

- Biotechnologies and sustainable development are often considered as **antagonists**, especially in Europe where **GMOs** (Genetically Modified Organisms) and agricultural biotechnology are opposed to more **environment-friendly** agricultural practices.
- Biotechnologies are often considered as tool for “agro-food giants” and hyper-intensive agriculture for which high investments must be balanced by high returns.
- This type of **high-tech agriculture** is opposed to **organic farming** and its role for the agriculture of developing countries is questioned



Sustainable agriculture must rely on solid and updated biological basis

- **Ecological intensification** is based on the understanding of biological functions in order to better exploit natural living resources while preserving them for future generations.
- **Biotechnologies** are based on mechanisms that nature already uses. When applied in appropriate ways the results can offer new solutions to existing problems.
- **Modern biology** is now integrating systems biology with very close links with biotechnology.

The basics of system biology

All living organisms are supported by the genome (DNA) through the action of the transcriptome (RNAs), proteome, metabolome, and ionome, the four basic pillars of functional genomics

RNAs

Coding and non coding

TRANSCRIPTOME

PROTEINS

Structure, enzymes

PROTEOME

GLUCIDS

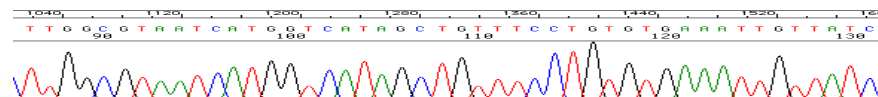
LIPIDS

METABOLOME

MINERALS

IONOME

GENOME

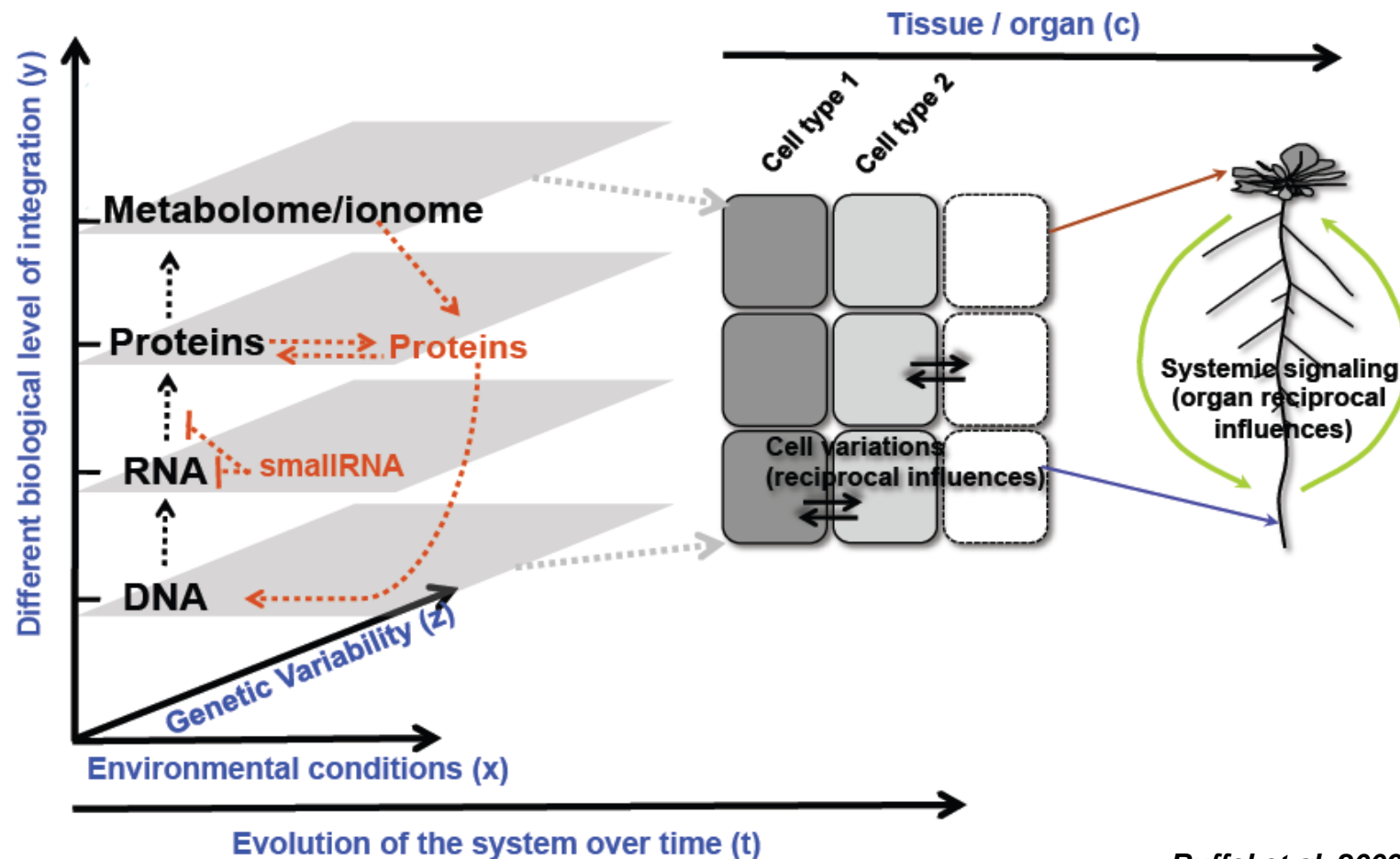


A new vision for plant physiology

- In order to study the **dynamics of plant metabolism** under stress and unravel regulatory mechanisms in place, it is important to combine the traditional more descriptive physiological approaches with the techniques of **functional genomics**, namely the high throughput methods for transcriptomic, proteomic, metabolomic and ionomic analysis.
- Using this **integrated analysis** it would be possible to study the dynamics of plant metabolism in the context of the plant system as a whole

(Chaves et al, 2009).

Plant systems biology: A dimension-wise challenge



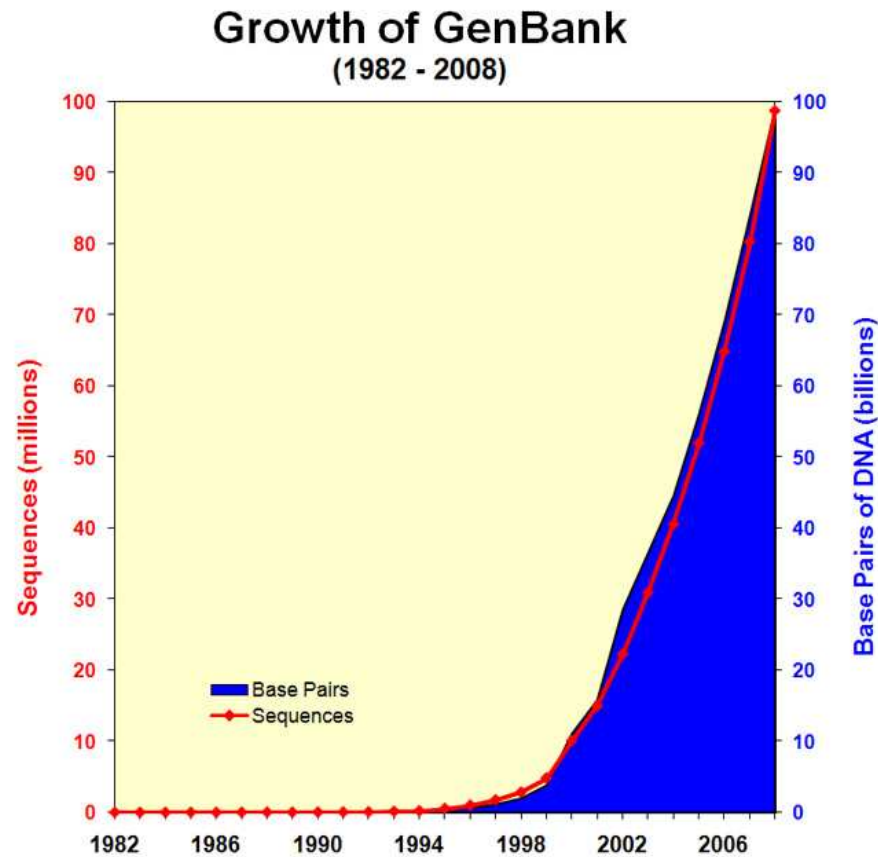
Ruffel et al, 2009. *Plant Physiology*

Amazing progress in Plant Genomics

- Automated and high-throughput technologies are quicker and cheaper: a few months for a new rice genome...
- To date, 1,193 genomes have been fully sequenced, including 77 archaeal, 996 bacterial and 123 eukaryal genomes.
- Ten sequenced and accessible plant genomes: *Cucumis*, *Vitis*, *Populus*, *Oryza*, *Arabidopsis*, *Glycine*, *Physcomitrella*, *Sorghum*, *Vitis* and *Zea* ... and much more under way: cassava, tomato, potato, papaya, cotton, eucalyptus, citrus, barley ...

Sources: GOLD, 2010; NCBI, 2010.

Huge amounts of genomics information is made available through databanks



Deciphering the oil palm genome has become reality

Zieler et al (2010) very recently reported on the generation whole-genome shotgun sequences of the oil palm genome. This breakthrough was achieved by a consortium joining **Synthetic Genomics Inc. (USA)**, **ACGT Sdn. Bhd. (Malaysia)** and the **J. Craig Venter Institute (USA)**.

In November 2009, The Malaysian Palm Oil Board (MPOB) announced that a consortium co-led by the Advanced Biotechnology and Breeding Centre has sequenced three oil palm genomes from the two species *Elaies oleifera* and *Elaeis guineensis*, including the *pisifera* and *dura* palms. **The consortium included St. Louis, Missouri-based Orion Genomics, MOgene LC and The Genome Center at Washington University, South Korea-based Macrogen Inc., and Adelaide, Australia-based GeneWorks Pty Ltd. (Wahid, 2009).**

In May 2009, a Malaysian consortium created by **Sime Darby Bhd**, one of the world's largest oil palm company and **Synamatix Sdn Bhd**, a bio-informatics company had announced the successful sequencing, assembling and annotating of the oil palm genome with 93.8% completeness.

Next Generation Sequencing (NGS)

Four different NGS technologies are now commercially available:

- 454 Life Sciences (acquired by Roche)
- Solexa (acquired by Illumina),
- ABI SOLID (acquired from Agencourt Biosciences), and
- Helicos Biosciences.

They differ in the length and quantity of generated sequences

- Solexa, ABI SOLID, and Helicos all produce very short reads in very large quantities,
- 454 platform can produce a more moderate amount of sequence, but with much longer read lengths.

Biotechnologies already impacts oil palm breeding

- **Potential yield** of the crop is 17 t of oil per hectare (Corley 1983).
- **Genetic variation** still encountered suggest that there exists considerable scope for improvement in yield, disease resistance and oil composition
- **Micropropagation** has been the first biotechnological application to impact oil palm breeding:
 - Several millions of clonal oil palms have been produced and are now field-monitored
 - Research now focus on
 - Reducing production costs
 - Early detection of *mantled* variants

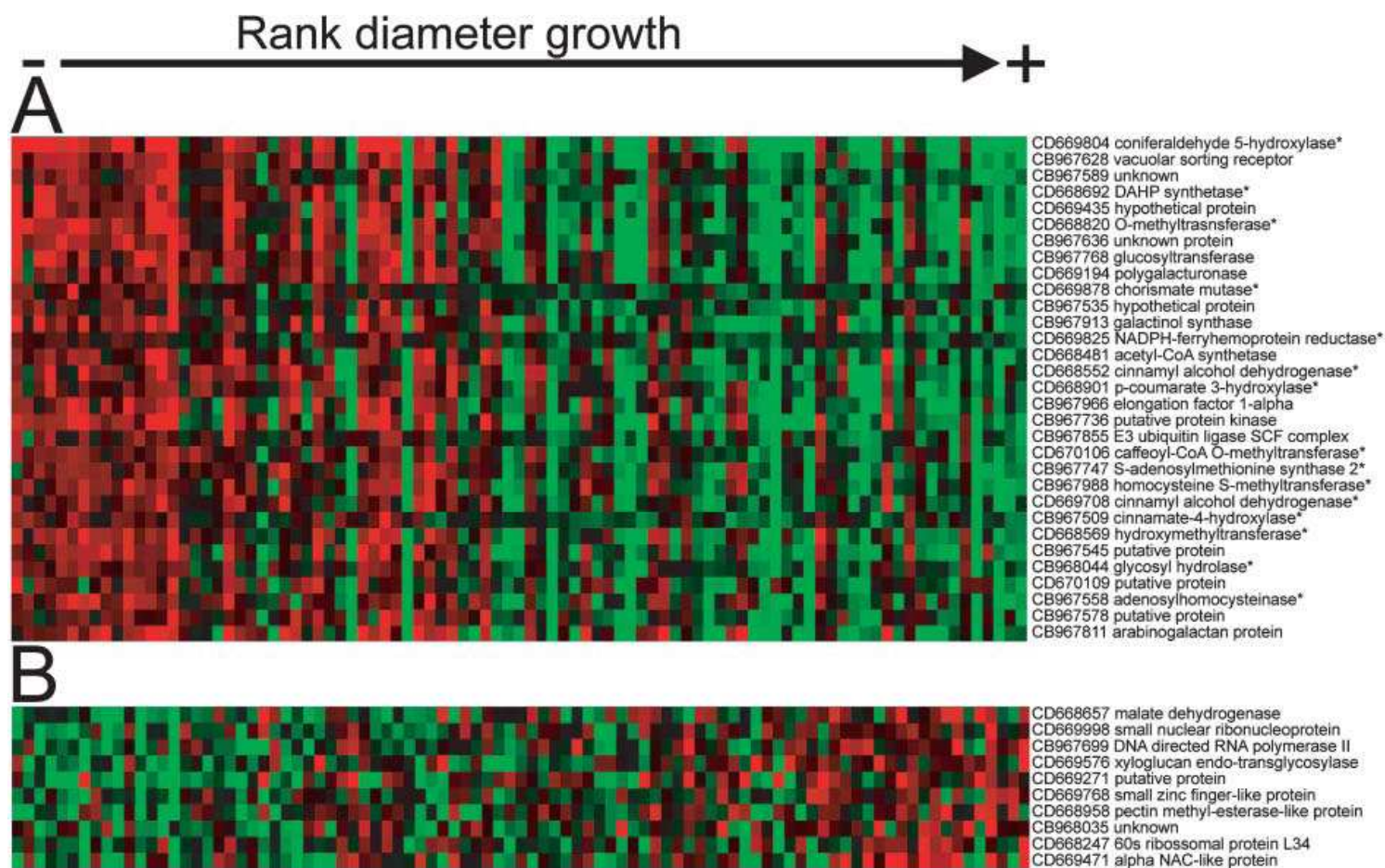
Molecular Breeding is under way

- **Molecular markers** help in selecting within material earlier in the breeding cycle to reduce the generational times
- **Fingerprinting and Linkage Studies** make use of RFLPs, AFLPs and Microsatelites Markers
- **Marker-assisted selection (MAS)** would allow the breeding process to be speeded up.
 - A marker for shell thickness potentially could have a high commercial value in breeding programs.
 - Determining whether selected palms are *pisifera* or sterile *tenera* before progeny testing.
 - Assessing the purity of commercial *tenera* seed lots.

Potential Applications of Molecular Markers

- Association of markers and phenotypic characters and permit **quantitative trait locus (QTL) analysis**.
- **Specifically designed field trials** are being exploited for segregation studies and QTL analysis for complex traits such as oil yield, oil/bunch and its components and vegetative characters.
- The recent development of the **CIRAD genetic map** will allow a detailed QTL analysis as a QTL for palm height (LINK2PALM) has already been identified.
- **The best is yet to come...** QTLs for fusariosis, bud rot tolerance, etc

Gene discovery through transcriptome analysis



eQTLs: Linking Transcriptomics with genetic maps

- **Microarray (DNA chips)** analysis approaches for transcriptomics allows an entire transcriptome to be analysed at a specific developmental point, under a particular stimulus, or in a particular tissue.
- Integrating **quantitative analysis of gene expression data** and **genetic map information** to discover genetic and metabolic networks regulating complex biological traits.
- Such resources are now being generated at a large scale in oil palm as part of an industrial consortium called OPGP for Oil Palm Genome Project.

Oil Palm genetic engineering: a valuable tool for functional genomics

- The initial transformation of oil palm with a marker gene was been reported by Parveez et al (2000)
- Genetic engineering in oil palm is progressing at slow pace
- Development is hampered by:
 - Public acceptance (palm oil being the only 100% GMO-free vegetable oil)
 - Biosafety issues (where to plant ?)
 - Genetic complexity of desired traits (from the production of bio plastics to the introduction of resistance to major pests and diseases)

The basics of system biology

All living organisms are supported by the genome (DNA) through the action of the transcriptome (RNAs), proteome, metabolome, and ionome, the four basic pillars of functional genomics

RNAs

Coding and non coding

TRANSCRIPTOME

PROTEINS

Structure, enzymes

PROTEOME

GLUCIDS

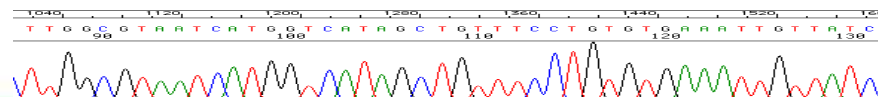
LIPIDS

METABOLOME

MINERALS

IONOME

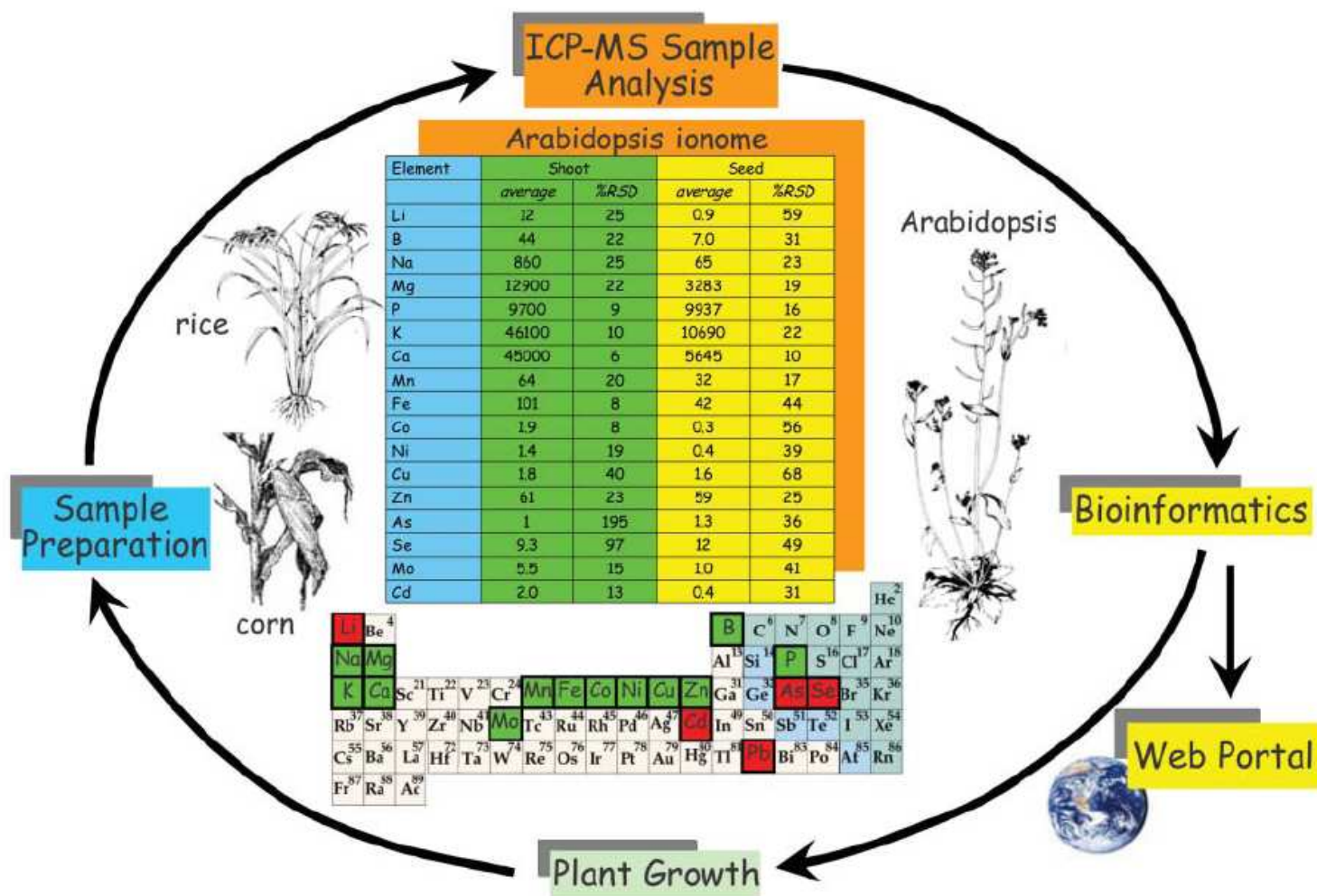
GENOME



Plant Ionomics: studying the social network of mineral nutrients

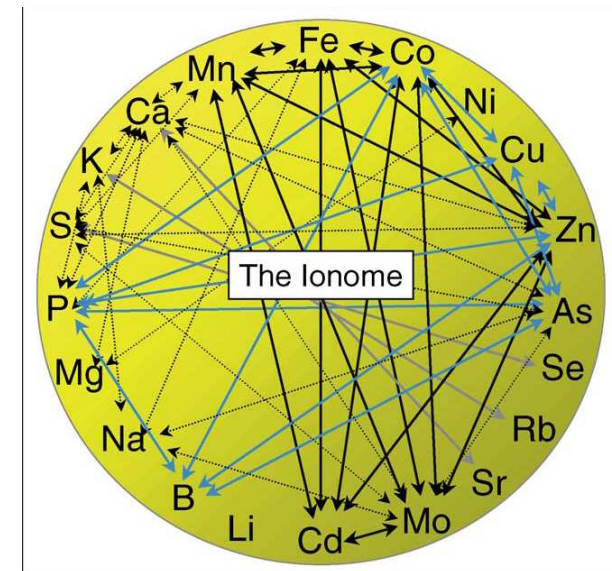
- The accumulation of a given element is a **complex process**, controlled by a **network of gene products** critical for uptake, binding, transportation, and sequestration.
- Many of these genes and physiological processes affect more than one element. It is necessary to **measure as many of the elements** contained in a cell, tissue, or organism (the ionome) as possible.
- Today, **high-throughput elemental profiling studies** are able to show **how the ionome responds to the environment** or to explore the **genetics that control the ionome**.
- When combined with new genotyping technologies, ionomics provides a rapid way to **identify genes that control elemental accumulation in plants**.

Plant ionomics: a promising area for oil palm



Ionomics imaging: getting closer ...

- Use of laser ablation sampling coupled to ICP-MS holds promise for development of high resolution ionomic imaging in living plant tissue
- Such technology would open up a completely new window onto the ionome, allowing changes in the total shoot, seed, or root ionomes, for example, to be mapped to specific tissues and cell types.
- Such ionomic imaging would also allow colocalization of in vivo gene expression and protein localization patterns with ionomic changes, providing spatial linkage between gene, protein, and ionomic function.



The indispensable cross-talk between Biology and Informatics

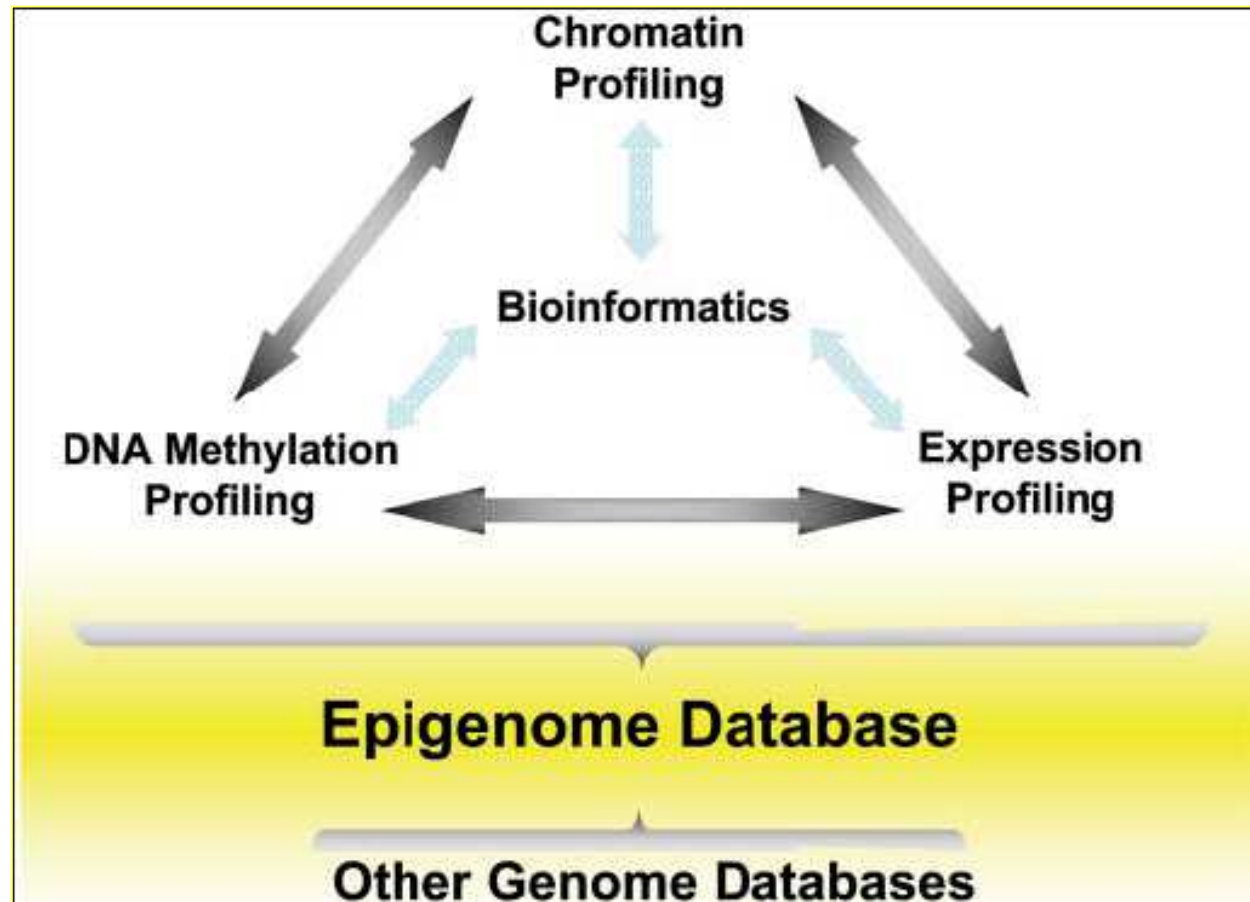
- ✓ Newest high-throughput technologies are now able to provide plant breeders and molecular physiologists with an **unprecedented amount of data**
- ✓ The treatment of these data is becoming **a scientific area by itself**, generating the use of complex algorithms, the management and safe exchange of data, computerized pipelines, etc...
- ✓ The challenge for biologists is to keep in mind the **real biological questions to be answered ...**

The emergence of Bioinformatics

Bioinformatics is a critical resource for the future of plant research and plant breeding. Current purposes to bioinformatics, in the context of plant and breeding research, are:

- 1) **To act as repository of information** generated with links to the associated physical resources
- 2) To allow the **repository to be searched** for specific resources or information.
- 3) To **develop linkages** between different types of information (sequences, map positions, alleles within a population, related genes) and develop tools to facilitate discovery of further links (gene prediction, protein interactions etc.)

Epigenetics: beyond the double helix



Murell et al, Human Molecular Genetics 2005

One genome, many epigenomes...

The **'epigenetic code'** considerably extends the information potential of the genetic code.

Thus, **one genome can generate many 'epigenomes'** as the fertilised egg progresses through development and translates its information into a multitude of cell fates.

The **epigenome is not static** and can be molded by developmental signals, environmental perturbations, and disease states.

Therefore, many epigenomes will need to be sequenced for a single organism, making epigenome sequencing perhaps even more challenging than genome sequencing.

Epigenetics and the oil palm: deciphering the « mantled » abnormality



- The commercial development of large scale propagation of oil palm is hampered by the generation of somaclonal variants affecting the floral architecture: **the « mantled » phenotype.**
- Field observation of clonal plantings on the long term and results from DNA markers analysis are consistent with an **epigenetic origin** for somaclonal variation.
- The “mantled” phenotype is correlated with a **global hypomethylation** of genomic DNA, which was revealed through a whole genome approach and the use of Methylation-Sensitive DNA markers, such as MS-RFLPS and MSAPs

An open access to genomics resources for the oil palm biotech community?

A key for the successful integration of biotechnologies in the improvement of oil palm sustainability

- A huge and increasing amount of data to manage and sort
- A real (and costly) risk of overlapping activities
- A multi-players game
 - ✓ Traditional: Oil Palm Companies (Sime Darby) , Institutions (MPOB, Cirad)
 - ✓ Newcomers: Bioinformatics, sequencing companies (DNA is DNA)
 - ✓ Consortia

Opening a public access to databases is essential

Profile



Alain RIVAL, PhD

Company / Institution

Cirad

Position

Oil Palm Research Coordinator

Contact

UMR DIAPC

**911, Avenue Agropolis
34394 Montpellier, France**

e-mail

Alain.rival@cirad.fr

Experience

- 1989** PhD Plant Biotechnology, University of Montpellier, France
- 1989** Research officer. Palms biotechnology, Cirad, Montpellier
- 1997** Research Director Habilitation, University of Paris, France
- 2000** Head of Oil Palm Programme, Cirad, Montpellier, France
- 2002** Qualified Professor in Biology
- 2004** Marie Curie European Fellow, CSIRO Plant Industry, Australia

[Draft and Slide Show]



International Conference on Oil Palm & Environment

Measurement and mitigation of environmental impact of palm oil production

23 – 25 February 2010 • Grand Hyatt Hotel, Nusa Dua Bali - Indonesia

Terima kasih !
Thank you very much!
Merci beaucoup !

Acknowledged by:



Organized By:



WWW.ICOPE-SERIES.com